7-2011

In High School and Pregnant: The Importance of Educational and Fertility Expectations for Subsequent Outcomes

Olga Yakusheva
Marquette University, olga.yakusheva@marquette.edu

Economic Inquiry, Vol. 49, No. 3 (July 2011): 810-837. DOI.

In High School and Pregnant: The Importance of Educational and Fertility Expectations for Subsequent Outcomes

Olga Yakusheva
Economics, Marquette University, Milwaukee, WI

Abstract

This study uses the High School and Beyond data (1980–1992) to examine the importance of educational and fertility expectations in explaining the achievement gap of adolescent mothers for over 5,500 young women from different socioeconomic backgrounds. Using a non-parametric local propensity score regression, the study finds that the economic disadvantage associated with having a child in high school is particularly large in poor socioeconomic environments; however, this disadvantage is a result of preexisting differences in the educational and fertility expectations and is not because of a diminished capacity of the socioeconomic environment to mediate the effect of an unplanned childbirth. The findings suggest that childcare assistance and other policies designed to alleviate the burden of child rearing for young mothers of low means may not produce the desired

improvement in their subsequent educational and labor market outcomes. A much earlier policy intervention with a focus on fostering young women's outlook for the future is needed.

Introduction

After an almost 40% decline in the incidence of teenage childbearing in the United States over a period of 15 consecutive years, the number of children born to adolescent mothers rose for the first time in 2005 ([15]). With a number of high profile celebrity teenage pregnancies ([13]), and a recent widely publicized case of a high school pregnancy pact ([21]), our nation's attention is once again drawn to the issue of adolescent childbirth and its impact on the young mother's future. This study finds that although the negative effect of a high school pregnancy on subsequent educational and earnings outcomes appears particularly large for young women from poor socioeconomic environments, it can almost fully be explained by preexisting low educational expectations and early fertility expectations, rather than by difficulties in coping with unplanned parenthood.

Over three decades ago, researchers first reported a large negative effect of early childbearing on the educational attainment of women ([31]). In a wealth of subsequent studies, adolescent mothers were found to compare unfavorably to women who delay childbearing along a wide range of measures of economic well-being, from high school completion, to postsecondary educational attainment, to labor market employment and earnings. However, the difficulty with interpreting these associations between early childbirth and poor subsequent outcomes lies in the fact that the timing of childbearing is clearly nonrandom. In particular, women from poor socioeconomic backgrounds are likely to have low educational expectations and low expected labor market earnings ([29]; [30]); thus, facing a relatively low opportunity cost of childbearing, they may choose to have a child earlier or to be more likely to engage in behaviors that may lead to an unwanted pregnancy ([25]). Women with high educational expectations and a high opportunity cost of childbirth, on the other hand, may be less likely to become pregnant; furthermore, given that 40% of unwanted pregnancies in the United States are terminated ([11]), these women may be more likely to terminate a pregnancy if one were to occur.

As educational and fertility expectations are private information that is not directly observable to the researcher, previous studies have employed a number of ways of accounting for the endogeneity of fertility timing, such as controlling for observable socioeconomic and demographic characteristics, using the siblings of adolescent mothers as the comparison group, and exploiting natural experiments like twin births and miscarriages (see [17] for a review of the literature). [17] carried out arguably the most convincing study of the causal impact of teenage childbearing. In their study, they showed that long-term outcomes of women whose teenage pregnancy ends in a miscarriage, instead of a live birth, are nevertheless very similar to outcomes of teenage mothers, thus suggesting that teenage childbirth itself has a few long-term negative consequences.

In this article, we come to a similar conclusion by using the educational and fertility expectation data for young women from the Sophomore Cohort of the High School and Beyond (HS&B) study (1980–1992) to *explicitly* account for the endogeneity of fertility timing. This approach is motivated by economic theory. If young women behave optimally and form rational expectations over their future fertility, education, and labor market outcomes, then measuring these expectations prior to childbirth should capture all relevant unobservable private information. Consequently, examining how fertility and educational expectations differ between adolescent mothers and women who delay childbearing can help explain why [17] did not find any *causal* effect of childbirth on subsequent educational and labor market outcomes despite the wealth of statistical evidence of the relationship.

Although researchers in the past questioned consistency and predictive validity of subjective expectations ([1]; [19]), a body of recent research on subjective expectations finds that individuals are fully capable of providing meaningful and informative answers to questions eliciting probabilistic expectations about significant future

events (see [23] for a review of the literature). Researchers also showed that meaningful responses to expectations questions can be elicited from minors, even as young as high school sophomores ([9]). Manski's (2004) findings gave rise to a slew of recent studies of subjective expectations in various areas of economics including migration ([26]), health ([8]), and education ([2]; [18]; Kaufmann 2009; [27]).

Simple tabulations of our study's expectations and outcomes data, shown in Table 1, yield two important observations. The first observation is that, consistent with the theory of optimal behavior, a woman's expected age of first childbirth is strongly and positively associated with her expected level of schooling. Among women with low educational expectations (no more than high school, second column in Table 1), the average expected age at first childbirth is 19.3, compared to 23.2 among women with high educational expectations (a college degree, last column in Table 1). The second observation is that there is a strong correspondence between subjective expectations and subsequent outcomes, suggesting that many women in our sample were able to form rational expectations over their future educational and fertility outcomes, and to follow through with their expectations. For example, the proportion of women who continued education beyond high school increases from less than 20% to over 70% as the expected level of schooling goes up, and the proportion of women who finished college increases from just over 2.5% to more than 55%. Women with high educational expectations also enter into parenthood at a later age and have fewer children. The link between educational and fertility expectations and between expectations and subsequent outcomes is the motivation for this study, as we explore the role of expectations in driving the gap in educational achievement and earnings between women who became pregnant in high school and women who had a child at an older age.

The remainder of the article is divided into four sections. Section II describes the data used in this article; Section III lays out the methodological approach; Section IV presents and discusses the results; robustness analyses are discussed in Section V; and finally, Section VI summarizes the findings and offers policy recommendations and directions for future research.

Table 1 Expectations and Outcomes

	Educational Expectations (1980)				
	All (All = 5,540; P[a] = 290; NP[b] = 5,250)	**School Degree Only (All = 1,100; P = 90; NP = 1,010)**	**Less Than 4-Year College (All = 1,620; P = 100; NP = 1,520)**	**4-Year College (All = 1260; P = 40; NP = 1,220)**	**More Than 4-Year College (All = 1110; P = 20; NP = 1,090)**
Fertility expectations (1980)					
Expected age at first childbirth, years	21.23	19.3	21.4	22.6	23.2
Fertility outcomes (1992)					
Age at first childbirth, years	25.2	21.39	23.14	25.01	26.35
Number of children	0.92	1.86	1.52	1.04	0.64
Educational outcomes (1982)					
Dropped out of high school					
All	5.34%	16.96%	3.34%	0.95%	0.80%
Pregnant in high school	15.38	30	8.16	0	8
Not pregnant in high school	4.8	15.81	3.02	0.98	0.64
Dropped out or delayed graduation					

All	12.21	32.21	9.21	3.89	4.22
Pregnant in high school	27.62	51.11	16.32	5.26	2
Not pregnant in high school	11.37	30.53	8.75	3.85	3.86
Educational outcomes (1992)					
School degree or equivalent					
All	95.1	84.21	97.21	99.04	99.19
Pregnant in high school	84.95	71.11	91.83	100	92
Not pregnant in high school	95.56	85.37	97.56	99.01	99.35
Two-year college degree or higher					
All	49.69	19.41	46.19	64.62	70.08
Pregnant in high school	28.32	18.88	34.69	44.73	32
Not pregnant in high school	50.85	19.46	46.93	65.24	70.95
Four-year college degree or higher					
All	28.55	2.54	15.76	46.82	55.52
Pregnant in high school	5.94	1.11	7.14	13.15	12
Not pregnant in high school	29.79	2.66	16.32	47.86	56.52
Labor market outcomes (1992)					
Annual earnings (X1,000)					
All	$17.14	$12.26	$15.99	$20.11	$20.81
Pregnant in high school	12.69	9.88	13.43	17.8	14.82
Not pregnant in high school	17.39	12.47	16.16	20.18	20.94

Notes: 450 women did not have data about educational expectations. In compliance with the National Center for Educational Statistics restricted data rounding rules, all sample size numbers are rounded to the nearest 10.
[a]Pregnant in high school (had a child or were 3 or more months pregnant by May of 1982).
[b]Not pregnant in high school.

II. DATA

The study sample is obtained from the restricted use HS&B survey data. The survey was conducted by the U.S. Department of Education and it contains longitudinal data on a diversified representative sample of almost 15,000 sophomore high school students in 1980. Follow-up surveys were conducted in 1982, when the majority of the respondents finished high school, and later in 1984, 1986, and 1992. The survey contains a wide range of data on the students' family, school, and community characteristics, high school experiences, attitudes, and expectations, as well as subsequent educational, fertility, and labor market outcomes. The survey contains approximately 7,480[1] women, of which a little less than 2,040 were excluded from the analysis because they did not have necessary information to complete this study. (About one-half of the excluded observations were lost as a result of attrition, and the remaining excluded observations did not have data on the number and ages of children, educational attainment, or earnings.) The final study sample includes approximately 5,540 women.

The first distinct feature of the HS&B data is that they contain detailed information about the women's expectations of future educational, fertility, and labor market outcomes in 1980, when the women were still sophomores in high school. The survey includes questions about the expected level of schooling, expected age of getting married and having a child, occupational aspirations, expected age of having a place of your own, and many others. As previously noted, a woman's expectations of her future fertility outcomes, educational achievement, and employment and earnings, are highly correlated. We were able to identify two questions that explained .50–.92 of the variance in most other expectations responses. These two questions are the expected age at first childbirth (which includes a "do not expect to have children" response), and the highest expected

level of education. These two questions formed the basis for the expectations variables included in our analysis: "Does not plan to have children," a dichotomous variable equal to one if a woman said she did not expect to have children, "Expected age at 1st childbirth," a continuous variable equal to the age a woman said she was expecting to have her first child, and three dichotomous variables for the expected schooling level, "two-year college degree/vocational degree/trade degree/certificate," "four-year college degree," and "advanced degree" (reference category is a high school diploma or less). The variables are described in Table 2.

Table 2 Variables and Descriptive Statistics

	Number of Observations	**Mean**	**Standard Deviation**	**Minimum Value**	**Maximum Value**
High school pregnancy (=1 if had a child or was at least 3-months pregnant in June 1982; =0 otherwise)	5,540	0.05	0.21	0	1
Outcomes					
High school Dropout, 1982 (=1 left high school without a diploma; =0 otherwise)	5,540	0.05	0.20	0	1
High school Dropout or Delay, 1982 (=1 if left high school without a diploma or if high school diploma awarded after June 1982; =0 otherwise)	5,540	0.12	0.31	0	1
High school degree, 1992 (=1 if high school diploma or GED by 1992; =0 if otherwise)	5,540	0.95	0.19	0	1
Two-year degree or higher, 1992 (=1 if has an Associate's degree, a Certificate or a Vocational/Trade school degree, a Bachelor's degree, or higher by 1992; =0 if otherwise)	5,540	0.49	0.41	0	1
Four-year degree or higher, 1992 (=1 if has a Bachelor's degree or higher by 1992; =0 if otherwise)	5,540	0.28	0.50	0	1
Earnings, 1992 (yearly income from wages and salaries in 1991, $1,000)	5,540	17.14	13.87	0	255
Individual characteristics					
Parental family income, 1980 ($1,000, top coded)	4,250	18.81	9.42	3.50	38.00
Father has a high school diploma, 1980 (0/1)	5,090	0.23	0.42	0	1
Father has some college, 1980 (0/1)	5,090	0.31	0.46	0	1
Single-parent household, 1980 (0/1)	4,950	0.19	0.40	0	1
Number of siblings, 1980	4,950	2.95	1.76	0	6
Religious, 1980 (scale 1–3; 1 = not at all, 3 = very much)	5,540	1.94	0.57	1	3
Black (0/1)	5,540	0.13	0.33	0	1
Hispanic (0/1)	4,890	0.20	0.40	0	1
Community characteristics					
% minority in student population, 1980 (0–100)	4,930	13.61	23.19	0	100
SMSA per capita personal income, 1980 ($1,000)	5,540	9.43	1.59	4.87	13.92
Expectations					

Plans to have children, 1980 (0/1)	5,540	0.90	0.30	0	1
Age expects to have a child, 1980	5,080	20.29	3.84	15	30
Expects 2-year college degree: expects to get less 2 years of college, vocational or trade school, 1980 (0/1)	5,090	0.31	0.46	0	1
Expects 4-year college degree, 1980 (0/1)	5,090	0.24	0.43	0	1
Expects to get an advanced degree, 1980 (0/1)	5,090	0.20	0.40	0	1
Dating frequency, 1980 (scale 1–4; 1 = never, 4 = almost every day)	5,540	2.03	0.97	1	4
Has a steady relationship, 1980 (0/1)	5,540	0.43	0.49	0	1
Math Test Score, 1980 (continuous variable, contains a composite of individual's scores on math item response test questions)	4,840	12.87	9.36	−4.48	37.99
Vocabulary Test Score, 1980 (continuous variable, contains a composite of individual's scores on vocabulary item response test questions)	4,390	8.80	5.33	0.90	21.00

Notes: 5,540 observations are included in all subsequent regressions. Missing values are replaced with zeros and missing 0/1 indicators are included in all models (but not reported). Other variables included in all models (but not reported) are eight regional indicators. In compliance with the National Center for Educational Statistics restricted data rounding rules, all sample size numbers are rounded to the nearest 10.

While previous research found subjective expectations to be meaningful and informative, the accuracy of measured expectations can depend on the wording of the questions. In particular, [23] calls for extra caution when, as it is the case in HS&B, expectations are elicited as verbal responses to questions about real-valued variables (e.g., "what is your expected age at first childbirth?"), as opposed to probabilistic responses to questions about binary outcomes (e.g., "what are the prospects that you have at least one child before the age of 20?"). Recognizing that responses to the HS&B survey questions could be a noisy measure of the underlying expectations, we add measures of behaviors that are likely to *reveal* educational and fertility expectations, like interactions with the opposite sex, measured by frequency of dating and whether or not a woman is involved in a steady relationship, and academic performance, measured by math and vocabulary test scores. These variables are also taken from the first wave of the survey administered in 1980 during the sophomore year of high school.

In addition to questions soliciting expectations, a second distinct feature of the HS&B restricted use data is that the 1992 follow-up has information on the ages of the women's children, in months, as well as the exact date that each of the women left high school (either graduating with a diploma or dropping out without earning a high school diploma). This information was used to create the "High school pregnancy" variable as an indicator equal to one if a woman had a child or was at least 3-months pregnant in high school. Almost 290 young women in the study sample reported having at least one child old enough to have been conceived while the mother was still in high school; a little more than 250 of these women had a child in high school, and 30 women were still pregnant at the time they left high school. Women who became pregnant after leaving high school were coded as not having experienced a high school pregnancy. In contrast to many earlier studies that defined early childbirth as having a child prior to achieving a certain age, usually during the teenage years ([12]; [17]; [16]) this study has an important advantage. Knowing the exact date of leaving high school and the ages of children allows us to exclude young women who became pregnant subsequent to dropping out of high school. While this does not eliminate all simultaneity between childbearing and high school graduation, the results of our study are likely to reflect the impact of a high school pregnancy on high school completion more accurately.

We examine both proximate and distal outcomes. The proximate outcomes include (1) dropping out of high school (a dichotomous variable equal to one if left high school without a diploma); and (2) dropping out or delaying high school graduation (equal to one if left high school without a diploma or received a high school diploma after July of 1982). The distal outcomes are measured in 1992 when the women were in their late twenties, and include (1) having a high school diploma or a GED; (2) having a 2-year postsecondary degree or higher; (3) having a 4-year college degree or higher; and (4) earnings. As the survey does not provide information on hours worked or part-time employment status after 1986, yearly earnings in 1991[2] are used as the only labor market outcome variable.

In addition to the expectations variables, we use an extensive set of socioeconomic and demographic characteristics of the women in 1980: parental family income and education, parental family structure, religiosity, as well as school and community characteristics (Table 2 contains the full set of variables included in the analysis). All of these variables have been linked to early childbearing in the literature ([3]; [4]; [6]; [22]). It should be noted that as the sample consists of high school sophomores in 1980, variation in ages is trivial and is not controlled for in the analyses. Also, many of the predictor variables have missing values. Observations with more than five missing variables were removed from the sample. In the remaining cases, missing values were substituted with zeros and missing value indicators were included in all regressions.

III. METHODS

To examine the role of educational and fertility expectations in creating the gap in educational achievement and earnings between young mothers and women who delayed childbirth, this study uses a locally weighted propensity score regression ([28]). The method allows us to identify women with similar a priori likelihoods of having an adolescent pregnancy (called *propensity scores*), so that a meaningful comparison could be made between the outcomes of women who became mothers early, and those who did not. If the relationship between early childbearing and poor subsequent outcomes is driven in part by optimal decision making, accounting for educational and fertility expectations in the propensity scores will significantly reduce conditional gaps in educational and fertility outcomes between young mothers and women who delayed childbirth.

In fact, if educational and fertility expectations captured *all* unobserved heterogeneity (Rosenbaum and Rubin's 1983 Strongly Ignorable Treatment assumption), the pregnancy variable and the residuals in the outcomes regressions would be uncorrelated conditional on the propensity score, and the coefficient of the pregnancy variable would be an unbiased estimate of the causal effect of an early pregnancy. However, because of the fact that subjective expectations may not be a fully accurate reflection of an individual's reality and that our ability to measure them may be limited, this study does not assume that unobserved heterogeneity can be fully eliminated by conditioning on expectations. Rather, we explore whether or not subjective educational and fertility expectations capture at least a part of unobserved heterogeneity in the relationship between the timing of childbirth and subsequent outcomes.

This study uses a two-step locally weighted regression technique. In the first step, we estimate a propensity score model and obtain predicted likelihoods of becoming pregnant in high school, or propensity scores, for all women in our sample, regardless of their childbearing status. In the second step, we estimate the statistical effect of a high school pregnancy on subsequent educational and labor market outcomes by stratifying the sample on the propensity score variables and carrying out local non-parametric estimation at a set of points along the range of the propensity score ([10]). First we estimate the propensity model without controlling for expectations, and then repeat the two-step estimation procedure after including controls for educational and fertility expectations in the propensity score model. If expectations play an important role, we expect that coefficients obtained after including expectations in the first step are significantly lower in magnitude than the

coefficients that were obtained without accounting for expectations. A detailed description of the estimation procedure follows below.

Step 1: To estimate the likelihood of becoming pregnant in high school, we project the realized fertility outcomes of the women on a set of individual characteristics. Specifically, we estimate the following Probit models:

(1) $P(X_i = 1|\boldsymbol{Z}_i) = \boldsymbol{\Phi}(\alpha'\boldsymbol{Z}_i + u_i), i = 1, \dots, N$

(2) $P(X_i = 1|\boldsymbol{Z}_i, \boldsymbol{E}_i) = \boldsymbol{\Phi}\left(\alpha^{*'}\boldsymbol{Z}_i + \beta^{*'}\boldsymbol{E}_i + u_i^*\right), i = 1, \dots, N$

where X is the pregnancy variable, $\boldsymbol{Z}$ is a vector of individuals' socioeconomic and demographic characteristics, and $\boldsymbol{E}$ is a vector of educational and fertility expectations. $\boldsymbol{\Phi}$ is the cumulative density function of a standard normal distribution, α, α^*, and β are coefficients, and N is the sample size. The estimated coefficients are then used to predict two propensity scores $\hat{p}_i$ and $\hat{p}_i^*$, or the a priori likelihood of having a high school pregnancy for every woman in the sample,

(3) $\hat{p}_i = \boldsymbol{\Phi}(\hat{\alpha}'\boldsymbol{Z}_i), i = 1, \dots, N$

(4) $\hat{p}_i^* = \boldsymbol{\Phi}\left(\hat{\alpha}^{*\prime}\boldsymbol{Z}_i + \hat{\beta}^{*'}\boldsymbol{E}_i\right), i = 1, \dots, N$

Step 2: To estimate the statistical effect of childbirth on subsequent outcomes of women with similar a priori likelihoods of having a high school pregnancy, we use a locally weighted regression. Locally weighted regression performs estimation around a fixed value of the propensity by giving the most weight to the data points nearest to the point of estimation. To compute weights for the locally weighted regressions, we first rank women according to each of the propensity scores, $\hat{p}_i$ and $\hat{p}_i^*$, then we choose a five point equally spaced[3] estimation grid in each of the propensity score ranges. Finally, for each of the propensity score ranges, we create five sets of probability weights, according to [10]. Both sets of probability weights for a given value of the propensity are computed using the following formula:

(5) $w_i^k = 15/16(1 - ((\hat{p}_i - \hat{P}^k)/h^k)^2)^2 i = 1, \dots, N; k = 1, \dots, 5$

Here w_i^k, is the weight of observation i in the weight structure associated with estimation point $\hat{P}^k$, $\hat{p}_i$ is the predicted likelihood of pregnancy for observation i, and $h^k = \max\{[\hat{P}^k - \min(\hat{p}_i)], [\max(\hat{p}_i) - \hat{P}^k]\}$ is the width of the interval between estimation point $\hat{P}^k$ and either the maximum or the minimum boundary point of the propensity range. Weights for the model with expectations variables are similarly computed. Note that the weights are all less than one in absolute value and they are symmetrical around the estimation point. Within each set of weights, the estimation point is assigned the maximum weight and observations that are further away from it are assigned progressively lower weights. This weighting structure has higher efficiency than other conventional non-parametric regression methods and is easy to implement because it does not require modification for the boundary points ([10]).

To estimate the statistical effect of early pregnancy on a given outcome variable, we estimate a system of five weighted regressions of the outcome variable, Y, on the pregnancy variable, X, and the propensity score:[4]

(6) $P(Y_i = 1|X_i, \hat{p}_i) = \boldsymbol{\Phi}\left(\gamma_0^k + \gamma_1^k x_i + \gamma_2^k \hat{p}_i + \varepsilon_{ik}\right), k = 1, \dots, 5, i = 1, \dots, N$

(7) $P(Y_i = 1|X_i, \hat{p}_i) = \boldsymbol{\Phi}(\gamma_0^{*k} + \gamma_1^{*k} x_i + \gamma_2^{*k} + \hat{p}_i^* + \varepsilon_{ik}^*)\ k = 1, \dots, 5;\ i = 1, \dots, N$

The two systems, (6) and (7), are estimated using a seemingly unrelated estimation method ([32]), which allows for the error terms to be correlated between the five equations within each system, and computes robust variance–covariance matrices of the coefficients.[5]

Finally, for each outcome variable (i.e., educational achievement and earnings), the resulting series of five estimated coefficients of the childbirth variable, $\hat{\gamma}_1^k$, $k = 1,\ldots,5$ and $\hat{\gamma}_1^{*k}$, $k = 1,\ldots,5$ are used to test whether there is a significant association between high school pregnancy and the outcome variable locally at every estimation point, as well as globally. The global estimate is computed as a weighted average of the local estimates $\hat{\gamma}_1 = \sum_{k=1,\ldots,5} \delta^k \hat{\gamma}_1^k$, and $\gamma_1^* = \sum_{k=1,\ldots,5} \delta^{*k} \hat{\gamma}_1^{*k}$ where the weight of each local estimate, δ^k (and δ^{*k}) is determined based on the sum of local weights corresponding to each estimation point. We use standard hypothesis testing techniques but compute all test statistics using robust variance and covariance estimates of the local coefficients. We use two-tailed critical values.

The propensity score approach has both advantages and limitations. The first advantage of using propensity score methodology lies in the fact that it compares teenage mothers to young women who did not have a child in high school but are otherwise similar based on a set of observed characteristics that affect the risk of an adolescent pregnancy (e.g., parental income and education levels, single-parent households, community characteristics, etc.). This allows the study to focus on the relationship between expectations and subsequent fertility, educational, and labor market outcomes, while reducing the bias associated with heterogeneity in socioeconomic backgrounds. The second advantage of using propensity score methodology is that it offers a natural way of collapsing a multitude of dimensions of individual heterogeneity into a single dimension, an a priori risk of an adolescent childbirth. As educational and fertility expectations are strongly correlated with a young mother's socioeconomic background, examining the relationship between expectations, childbearing, and subsequent outcomes over a continuum of propensity scores provides a unique insight into how the relationship changes depending on the socioeconomic background of a young mother.

A limitation of using the propensity score approach is that the timing of childbearing may be determined by a large number of factors (both personal and environmental), many of which may be unobserved or poorly measured. Any residual simultaneity in fertility, and schooling, and labor market decisions resulting from such data limitations can result in a bias, or systematic overestimation of the negative relationship between early childbearing and subsequent outcomes. As the magnitude of this residual bias depends on the number and the quality of variables used to predict the propensity score ([14]), the propensity models in this study include a wide array of sociodemographic and personal characteristics, all of which come from a large, reliable, and nationally representative data source.[6] Nevertheless, we would like to emphasize that all coefficient estimates reported in this study should be interpreted as statistical associations between childbirth and subsequent schooling and labor market outcomes, rather than the causal effects of childbirth on subsequent outcomes.

IV. RESULTS

Table 3 shows the estimates of the propensity toward early pregnancy. First, the propensity is estimated using only socioeconomic and demographic characteristics, such as parental income and family structure, parental education, race, religion, and school and community characteristics. The second model estimates the propensity toward early childbearing by expanding the set of predictors to include measures of educational and fertility expectations. We also estimate a variant of the second model that allows for nonlinearities in the expectations variables and interaction terms between socioeconomic factors and expectations.

Table 3 Estimation of the Predicted Likelihood of Having a Child While in High School

	Dependent Variable = Had a Child Prior		

	to Leaving High School		
	1	**2**	**3**
Family income	−0.01 (2.36)**	−0.01 (1.39)	
Family income (7,000–12,000)			−0.85 (1.02)
Family income (12,000–16,000)			−1.29 (1.06)
Family income (16,000–20,000)			−1.28 (1.23)
Family income (20,000–25,000)			−2.39 (1.76)*
Family income (>25,000)			−1.52 (1.03)
Father has high school diploma	−0.05 (0.54)	−0.01 (0.06)	0.04 (0.41)
Father has some college	−0.25 (2.60)***	−0.06 (0.61)	−0.07 (0.64)
Number of siblings	0.11 (5.87)***	0.09 (4.60)***	−0.08 (3.64)***
Single-parent household	0.16 (2.20)**	0.15 (1.89)*	0.16 (1.85)*
Religious upbringing	−0.11 (1.88)*	−0.03 (0.46)	−0.06 (0.98)
Black	0.38 (3.72)***	0.36 (3.27)***	0.27 (2.07)**
Hispanic	0.10 (1.21)	0.01 (0.16)	−0.07 (0.70)
% Minority in student population	0.24 (1.70)*	0.18 (1.43)	0.24 (1.40)
SMSA per capita income	−0.06 (2.91)***	−0.05 (2.09)**	−0.03 (1.27)
Does not plan to have children		−0.75 (4.88)***	−2.93 (4.36)***
Expected age at first childbirth		−0.09 (6.70)***	−33.87 (3.04)***
Expected age at first childbirth, squared			2.11 (2.82)***
Expected age at first childbirth, cubed			−0.06 (2.63)***
Expected age at first childbirth, fourth degree			−0.00 (2.45)**
Expects 2-year college degree		−0.06 (0.60)	−0.14 (0.71)
Expects 4-year college degree		−0.12 (1.19)	−0.01 (0.03)
Expects advanced degree		−0.25 (2.05)**	−0.98 (2.30)**

Dating		0.06 (1.44)	−0.06 (0.75)
Steady relationship		0.16 (1.98)**	0.43 (2.25)**
Math test score		−0.19 (3.39)***	−0.23 (0.87)
Math test score, squared			0.00 (0.27)
Vocabulary test score		−0.01 (1.72)*	−0.05 (0.97)
Vocabulary test score, squared			0.00 (1.01)
Interaction effects	No	No	Yes
Pseudo R^2	0.09	0.16	0.22

Notes: Absolute value of *t* statistics in parentheses. Model 3 includes a full set of interaction effects between the family income indicators and each of the expectations variables.
*<10%. **<5%. ***<1%.

The estimates show that young women from higher income families with better educated parents are less likely to become pregnant in high school, as are women growing up in a two-parent household. African American women are at a higher risk of early motherhood. The incidence of early childbearing is higher in low-income communities and in schools with a higher proportion of minority students. The expectations variables, as a group, have a significant influence on the likelihood of having a high school pregnancy and including them in the propensity model increases the pseudo R^2 from .09 to.16. Having low educational expectations makes a young woman more likely to have a child in high school. Good academic performance, on the other hand, is associated with a lower risk of early childbearing. Not surprisingly, both dating and having a steady relationship are high-risk factors. Higher expected age at first childbirth is negatively related to the risk of a high school pregnancy, and, as suggested by the nonlinear specification of the relationship, the gradient increases very quickly after the expected age at first childbirth becomes greater than 20. Most of the interaction effects of the expectations variables with the income categories are statistically insignificant; however, the inclusion of nonlinear effects and interaction effects further increases the pseudo R^2 of the model from .16 to.22, and they are jointly statistically significant at a higher level than the.01 level. The distributions of the resulting propensity scores for the propensity model without the expectations measures and with the expectations measures (including nonlinear and interaction effects) are shown in Figure 1.

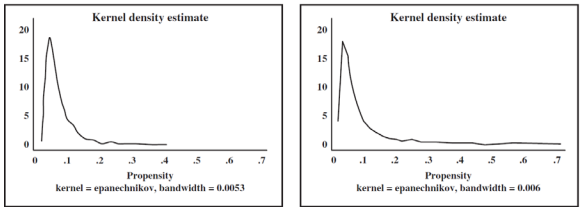

Figure: 1 Density Function of Propensity to Early Childbirth, without Controls for Expectations (Left) and with Controls for Expectations (Right)

The equally spaced 5-point estimation grid based on the propensity model without controls for expectations consists of the following points: (0,.09,.19,.28,.38); and the estimation grids based on the two propensity models with controls for expectations are the same to the second decimal point (0,.17,.34,.51,.67). Given limitations of regression analysis in predicting the risk of pregnancy (the pseudo R^2 of the propensity models is .09–.22), this

study treats propensity toward early childbearing as an ordinal, rather than a cardinal, concept. Therefore, in the remainder of the article the estimation points in all three models will be referred to as "low,""low-to-medium,""medium,""medium-to-high," and "high." Approximately 75% of women in the sample have low and low-to-medium propensities toward a high school pregnancy according to the propensity model without controls for expectations, and almost 85% of the women fall into the low and low-to-medium categories according to the propensity model with controls for expectations. Close to 5% and 8% of women have propensities in the high range according to the model with controls for expectations and the model without controls for expectations, respectively.

Estimation points associated with low propensities correspond to young women from better socioeconomic backgrounds. Prominent differences in socioeconomic characteristics across the estimation points can be easily seen in Table 4, which displays a series of locally weighted means of several variables describing a young woman's socioeconomic background. For example, average yearly parental family income at the low propensity estimation point is almost $15,000 (in 1980), compared to just under $9,000 at the high propensity estimation point. At the low propensity estimation point, almost 30% of women come from families with a college-educated father, compared to only 7% at the high estimation point. The average number of siblings increases from 2.52 to 3.81, as we move from low to high propensity; at the same time, the prevalence of single-parent families increases twofold, from less than 20–40%. Women with high propensities are more likely to be African American, and they attend schools with a higher proportion of minority students. Finally, high propensity is associated with living in a lower income community. Thus, locally weighted regression based on the propensity score provides a natural framework for examining the role of expectations in the relationship between an early pregnancy and subsequent educational and labor market outcomes along a continuum of socioeconomic environments.

Table 4 Means of Socioeconomic Characteristics by the Value of the Propensity Score

	Predicted Likelihood of High School Pregnancy				
	Low	**Low to Medium**	**Medium**	**Medium to High**	**High**
Family income	14.81	12.22	9.67	9.16	8.93
Father has high school diploma	0.20	0.20	0.18	0.17	0.16
Father has some college	0.29	0.27	0.09	0.07	0.07
Number of siblings	2.52	2.63	3.55	3.72	3.81
Single-parent household	0.18	0.20	0.35	0.39	0.40
Religious	1.74	1.72	1.59	1.57	1.56
African American	0.10	0.12	0.34	0.41	0.44
Hispanic	0.20	0.20	0.24	0.23	0.22
% Minority student	12.25	13.51	26.50	30.79	32.82
Per capita income	9.47	9.42	8.98	8.90	8.86

Notes: Weighted means of variables are given in all columns. The means were calculated using local non-parametric weights.

Table 5 shows the results of a locally weighted propensity regression of the effect of becoming pregnant in high school on various measures of educational achievement (columns 1–5), and on earnings (column 6), based on the propensity model without controls for expectations. The results based on the two propensity models that control for expectations are presented in Table 6 (without nonlinear and interaction effects) and Table 7 (with nonlinear and interaction effects). For each of the five propensity points listed in the first column, there are four rows of data showing the local marginal effect of the high school pregnancy variable, its robust *t*-statistic, and

the marginal effects and *t*-statistic of the propensity score variable. Finally, the weighted average of the local marginal effects and its chi-square statistic are displayed in the last two rows of the table.

Table 5 Non-Parametric Propensity Score Estimates, Not Controlling for Expectations in the Propensity Score Model

	Outcomes					
	1982		**1992**			
	High School Dropout	**High School Dropout/ Delay**	**High School Diploma**	**Two-Year Degree or Higher**	**Four-Year Degree or Higher**	**Earnings**
Local estimates						
At low risk						
High school pregnancy	0.032 (2.55)**	0.060 (3.01)***	−0.032 (2.64)***	−0.159 (4.39)***	−0.270 (7.40)***	−2.475 (2.60)***
Propensity	0.034 (14.53)***	0.078 (16.82)***	−0.031 (14.20)***	−0.223 (12.42)***	−0.390 (14.55)***	−5.972 (11.35)***
At low-to-medium risk						
High school pregnancy	0.075 (0.24)***	0.112 (3.67)***	−0.078 (3.26)***	−0.145 (4.89)***	−0.120 (9.30)***	−2.465 (2.70)***
Propensity	0.082 (6.85)***	0.140 (8.64)***	−0.078 (6.72)***	−0.193 (11.53)***	−0.206 (25.33)***	−5.428 (11.30)***
At medium risk						
High school pregnancy	0.098 (2.31)**	0.154 (3.20)***	−0.100 (2.39)**	−0.134 (3.16)***	−0.079 (4.61)***	−2.553 (2.24)**
Propensity	0.060 (2.48)**	0.074 (2.74)***	−0.054 (2.33)**	−0.046 (1.96)**	−0.062 (8.57)***	−2.587 (4.27)***
At medium-to-high risk						
High school pregnancy	0.129 (2.17)**	0.186 (3.08)***	−0.132 (2.26)**	−0.141 (2.77)***	−0.061 (2.70)***	−3.259 (2.37)**
Propensity	0.068 (1.98)**	0.063 (1.96)*	−0.064 (1.90)***	−0.036 (1.49)	−0.035 (9.18)***	−1.665 (2.56)**
At high risk						
High school pregnancy	0.156 (1.99)**	0.206 (2.79)***	−0.162 (2.07)**	−0.141 (2.29)**	−0.047 (1.59)	−3.572 (2.15)**
Propensity	0.078 (1.62)	0.06 (1.51)	−0.75 (1.56)**	−0.031 (1.20)	−0.23 (7.73)**	−1.295 (1.72)*
Weighted average of local estimates						
High school pregnancy	0.059 (15.98)***	0.094 (20.38)***	−0.061 (18.09)***	−0.151 (22.96)***	−0.186 (27.38)***	−2.522 (15.59)***

Notes: Shown are marginal effects from a Probit regression. Bottom row is weighted average of local estimates. Absolute values of *t* -statistic for local estimates and of chi-square statistic for the weighted average are in parenthesis.

*<10%. **<5%. ***<1%.

TABLE 6 Non-Parametric Propensity Score Estimates, Controlling for Expectations in the Propensity Score Model

	Outcomes					
	1982		**Outcomes**			

	High School Dropout	High School Dropout/ Delay	High School Diploma	Two-Year Degree or Higher	Four-Year Degree or Higher	Earnings
Local estimates						
At low risk						
High school pregnancy	0.004 (0.70)	0.005 (0.42)	−0.006 (0.98)	−0.073 (2.02)**	−0.199 (3.95)***	−1.065 (1.13)
Propensity	0.029 (15.04)***	0.075 (23.20)***	−0.026 (14.36)***	−0.249 (18.50)***	−0.620 (22.70)***	−5.058 (13.66)***
At low-to-medium risk						
High school pregnancy	0.024 (0.81)	0.023 (0.65)	−0.03 (1.17)	−0.049 (1.89)*	−0.006 (3.40)***	−1.004 (1.07)
Propensity	0.139 (8.62)***	0.21 (12.50)***	−0.126 (8.16)***	−0.128 (28.08)***	−0.028 (5.50)***	−4.524 (12.98)***
At medium risk						
High school pregnancy	0.053 (0.89)	0.085 (1.41)	−0.075 (1.30)	−0.034 (0.75)	−0.008 (1.18)	−1.326 (0.95)
Propensity	0.101 (3.36)***	0.127 (5.12)***	−0.082 (2.85)***	−0.038 (2.51)**	−0.011 (2.84)***	−1.288 (2.35)**
At medium-to-high risk						
High school pregnancy	0.026 (0.31)	0.061 (0.90)	−0.067 (0.81)	−0.037 (0.69)	−0.002 (0.55)	−1.248 (0.70)
Propensity	0.111 (3.37)***	0.106 (7.22)***	−0.09 (2.44)**	−0.029 (2.18)**	−0.002 (0.83)	−1.068 (1.67)*
At high risk						
High school pregnancy	0.008 (0.09)	0.038 (0.60)	−0.057 (0.54)	−0.033 (0.56)	−0.000 (0.28)	−1.206 (0.55)
Propensity	0.103 (6.78)***	0.074 (7.28)***	−0.093 (2.90)***	−0.023 (2.43)**	−0.000 (0.35)	−1.025 (1.35)
Weighted average of local estimates						
High school pregnancy	0.016 (0.78)	0.019 (0.69)	−0.023 (1.85)	−0.059 (3.04)*	−0.097 (6.83)***	−1.062 (2.63)

Notes: Shown are marginal effects from a Probit regression. Bottom row is weighted average of local estimates. Absolute values of t -statistic for local estimates and of chi-square statistic for the weighted average are in parenthesis.

*<10%. **<5%. ***<1%.

TABLE 7 Non-Parametric Propensity Score Estimates, Controlling for Expectations in the Propensity Score Model with Interaction Effects

	Outcomes					
	1982		1992			
	High School Dropout	High School Dropout/ Delay	High School Diploma	Two-Year Degree or Higher	Four-Year Degree or Higher	Earnings
Local estimates						
At low risk						

High school pregnancy	0.009 (1.06)	0.017 (1.09)	−0.011 (1.31)	−0.088 (2.37)**	−0.177 (3.83)***	−1.139 (1.17)
Propensity	0.029 (15.13)***	0.068 (19.28)***	−0.026 (14.44)***	−0.19 (14.68)***	−0.429 (18.41)***	−4.286 (12.32)***
At low-to-medium risk						
High school pregnancy	0.033 (1.20)	0.046 (1.35)	−0.041 (1.52)	−0.066 (2.27)**	−0.023 (4.01)***	−0.978 (1.03)
Propensity	0.099 (7.80)***	0.151 (10,322)***	−0.090 (7.40)***	−0.149 (17.65)***	−0.070 (11.54)***	−3.901 (12.02)***
At medium risk						
High school pregnancy	0.055 (1.06)	0.093 (1.64)	−0.065 (1.27)	−0.036 (0.76)	−0.025 (1.71)*	−0.414 (0.34)
Propensity	0.063 (2.59)***	0.077 (3.21)***	−0.059 (2.46)**	−0.037 (2.30)**	−0.018 (7.39)***	−1.482 (3.18)***
At medium-to-high risk						
High school pregnancy	0.034 (0.46)	0.066 (0.91)	−0.054 (0.75)	−0.035 (0.67)	−0.015 (0.99)	−0.310 (0.21)
Propensity	0.072 (2.39)**	0.080 (3.36)***	−0.065 (2.16)**	−0.029 (2.13)**	−0.008 (2.86)***	−1.132 (2.34)**
At high risk						
High school pregnancy	0.016 (0.17)	0.044 (0.54)	−0.042 (0.44)	−0.032 (0.54)	−0.007 (0.56)	−0.230 (0.13)
Propensity	0.079 (2.47)**	0.076 (4.50)***	−0.073 (2.11)**	−0.024 (2.24)**	−0.004 (1.06)	−1.067 (1.94)*
Weighted average of local estimates						
High school pregnancy	0.022 (1.83)	0.035 (2.36)	−0.028 (3.07)	−0.073 (4.78)*	−0.096 (8.75)***	−0.988 (2.66)

Notes: Shown are marginal effects from a Probit regression. Bottom row is weighted average of local estimates. Absolute values of t -statistic for local estimates and of chi-square statistic for the weighted average are in parenthesis.

*<10%. **<5%. ***<1%.

The results based on the propensity model without controls for expectations (Table 5) show that even after socioeconomic and demographic variables are accounted for in the propensity model, educational and labor market outcomes of women who had a child in high school are still consistently worse compared to women who delayed childbearing. Although the estimates of the average effects (at the bottom of the table) are smaller in magnitude than the unconditional differences that can be found by comparing the outcomes of young mothers and women who delayed childbirth in Table 1, the gaps remain prominent. Having a child or becoming pregnant in high school is associated with a 5.9% higher probability of dropping out of high school (compared to a 10.6% higher probability of dropping out in Table 1),[7] and many young mothers put off high school graduation. By their late twenties, women who became pregnant in high school are still 6.1% less likely to have a high school diploma (the unconditional difference in Table 1 is 10.6%), and they are 15.1% less likely to have postsecondary education (the unconditional difference in Table 1 is 22.5%). Interestingly, most of the gap in postsecondary educational achievement is because of a reduction in the likelihood of completing a 4-year college degree or a more advanced program. The reduction in the likelihood of having a 4-year college degree or higher is 18.6% (unconditional difference in Table 1 is 23.5%), which is larger in magnitude than the total reduction in the likelihood of having any postsecondary degree (15.1%), suggesting that the change in the likelihood of obtaining

a shorter term postsecondary degree after childbirth is small, and may even be positive. Finally, young mothers earn $2,522 less (roughly half of the unconditional difference) in yearly incomes compared to women who did not become pregnant in high school.

The propensity toward early childbirth is consistently, strongly, and negatively related to both immediate educational outcomes (increasing the risk of dropping out and delaying high school graduation) and educational outcomes later in life (reducing the likelihood of having a high school degree and postsecondary schooling); it is also negatively related to earnings.[8] This is not surprising as the socioeconomic and demographic factors embedded in the estimate of the a priori risk of a high school pregnancy are likely to have an impact on subsequent educational and labor market achievement, independently of their role as predictors of a high school pregnancy.

A closer examination of the series of local estimates reveals that the magnitudes of the pregnancy variable vary systematically with the propensity score. To help put the estimates in a context, the left-hand side panels of Figures 2–7 show the series of predicted local likelihoods based on estimated models in Table 5, for young mothers and for women who did not become pregnant in high school. The figures clearly show that as the predicted risk of an adolescent pregnancy increases, the educational and labor market outcomes deteriorate for both women who became pregnant in high school and women who did not. Furthermore, the educational achievement gap and the earnings gap of adolescent mothers become larger as the risk of an early pregnancy increases. For example, among women with low propensities (associated with better socioeconomic environments), adolescent mothers are only 3.2% less likely to have a high school diploma in 1992; however, the difference in predicted rates of high school graduation among women with high propensities is 16.2% (the local estimates can be found in Table 5). As the average rate of high school graduation is significantly lower among women with high propensities, the large absolute difference in the magnitude of the high school completion gap translates into an even larger relative gap in the likelihood of completing high school (Figure 4). The relative postsecondary educational achievement gap of young mothers (relative to women who did not have a child in high school) is significantly greater among women with high propensities than among women with low propensities (Figures 5 and 6). Finally, Figure 7 shows that the difference in earnings between young mothers and women who delayed childbearing also appears to be disproportionately large among women with high propensities than among women with lower predicted likelihoods of childbirth ($3,572, or a 32% reduction in earnings, compared to $2,475, or an 11% reduction in earnings, for high and low propensity women, respectively).

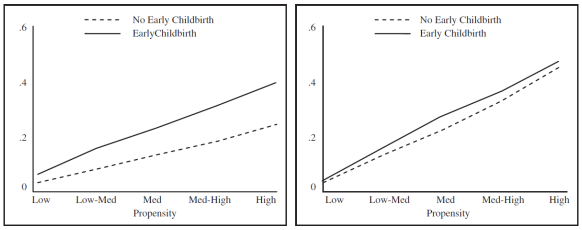

Figure: 2 Predicted Probability of Dropping Out of High School, without Controls for Expectations (Left) and with Controls for Expectations (Right)

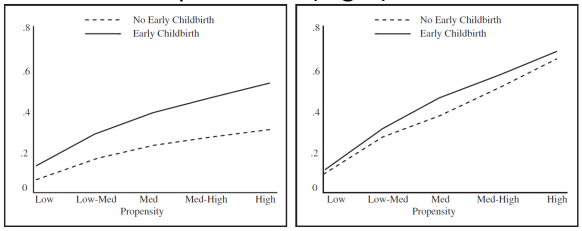

Figure: 3 Predicted Probability of Dropping Out of High School or Delaying Graduation, without Controls for Expectations (Left) and with Controls for Expectations (Right)

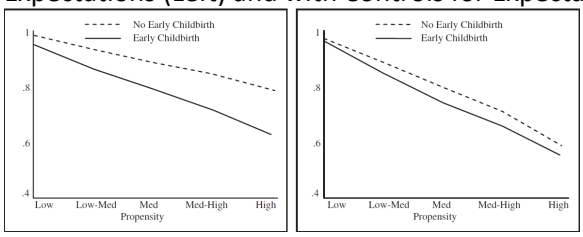

Figure: 4 Predicted Probability of Having a High School Degree by 1992, without Controls for Expectations (Left) and with Controls for Expectations (Right)

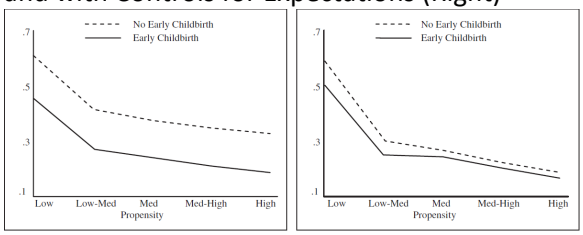

Figure: 5 Predicted Probability of Having a 2-Year College Degree or Higher by 1992, without Controls for Expectations (Left) and with Controls for Expectations (Right)

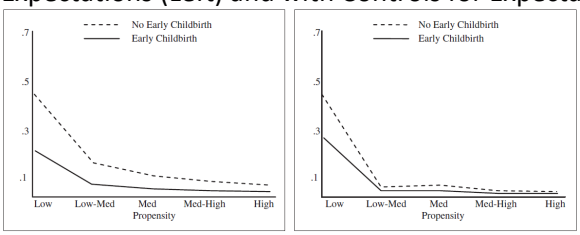

Figure: 6 Predicted Probability of Having a 4-Year College Degree or Higher by 1992, without Controls for Expectations (Left) and with Controls for Expectations (Right)

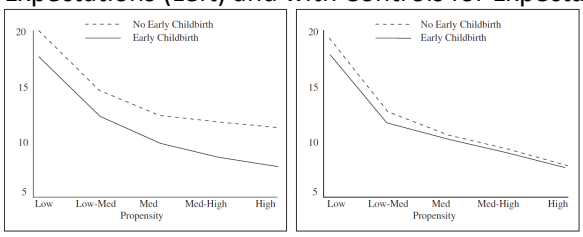

Figure: 7 Predicted Annual Earnings by 1992, without Controls for Expectations (Left) and with Controls for Expectations (Right)

These findings point to noticeable differences in the extent of the negative association between early childbirth and subsequent educational and labor market outcomes across socioeconomic groups. The observed disparities may be a result of a diminished capacity of young mothers from poor socioeconomic environments to cope with the demands of motherhood, possibly because of financial insecurity or a lack of parental support and guidance. Alternatively, the variation in the estimates may be a sign that factors not included in the first propensity model (i.e., educational and fertility expectations) play a more prominent role in shaping the relationship between childbirth, educational achievement, and labor market outcomes in poor socioeconomic environments.

The estimation results obtained after controlling for educational and fertility expectations in the propensity model are presented in Table 6 (based on the model without nonlinear or interaction effects) and Table 7 (based on the model with nonlinear and interaction effects). The series of local predicted values based on the estimates in Table 7 are graphed on the right-hand side panels of Figures 2–7. The results in the tables reveal that much of the association between having a child in high school and poor educational achievement and low earnings arises

because of differences in educational and fertility expectations that existed prior to childbirth. After controlling for the expectations variables, there are no statistically significant differences in the likelihoods of dropping out or delaying high school graduation, obtaining a high school diploma, or in earnings, between women who became pregnant in high school and those who did not.

Controlling for educational and fertility expectation reduces the average postsecondary achievement gap by roughly one-half of its magnitude in Table 5. The reduction in the magnitudes of the local estimates is also dramatic. With the exception of women with the lowest risk of pregnancy, the local estimates of the effect of childbirth on postsecondary achievement become very small in magnitude are no longer statistically significant. Figures 5 and 6 reveal that the postsecondary educational achievement gap decreases because controlling for educational and fertility expectations raises the predicted likelihood of having a postsecondary degree among women who became pregnant in high school, while lowering it among women who did not. The average effect in the model for a 2-year degree or above is only marginally significant $(p = .08)$, and as its magnitude continues to be smaller than in the model for a 4-year degree we can once again conclude that there is little negative (and possibly even positive) effect of a high school pregnancy on the likelihood of obtaining a 2-year postsecondary degree. Finally, after controlling for educational and fertility expectations, the reduction in labor market earnings associated with having a child in high school becomes less than half of the original magnitude (the average earnings gap is $1,062 and $988 in Tables 6 and 7, respectively, compared to $2,522 in Table 5), and it is no longer statistically significant overall or at any of the local estimation points.

The fact that most of the coefficient estimates in Tables 6 and 7 are statistically insignificant suggests that among women of the same socioeconomic status with similar educational and fertility expectations and academic performance, young mothers may not be at much of a disadvantage with respect to educational achievement and earnings. Although there is some evidence that for low risk women, becoming pregnant in high school appears to interfere with obtaining a 4-year college degree or higher, this result may have been driven by residual biases from omitted variables in the propensity models. It is also possible that the young age of the women (average age in 1992 is about 27 years old) precludes us from fully observing the life cycle consequences of an early childbirth. In particular, as this category includes postsecondary degrees that usually require 4–10 years of college, it is very likely that some young mothers may still be in the process of completing their degrees at the time of the last follow-up. Indeed, it has been shown in the previous literature that the negative effect of an early childbirth dissipates over time ([5]; [17]). Thus, our results are consistent with the recent instrumental variable estimates ([17]) that produce little evidence of a deleterious effect of an unplanned early pregnancy on life cycle academic achievement or labor market success.

Furthermore, the fact that the disproportionally large achievement gaps at the high values of propensity that we observed for educational and labor market outcomes in Table 5 become insignificant after controlling for educational and fertility expectations suggests that the reduction in schooling and earnings that follows an early childbirth among women of a low socioeconomic status is the result of their a priori lower educational and labor market aspirations. Thus, there is little evidence that the outcomes of low-income young mothers are further worsened by their diminished ability to stay the course after an unexpected early pregnancy or a lack of financial resources and social support to help them cope.

V. ROBUSTNESS ANALYSIS

We carried out three robustness checks. First, we modeled the propensity score variable in the second-stage regression as a linear spline rather than a single continuous variable. Second, we included the variables used to form the propensity scores into the second-stage model of the relationship between childbirth and outcomes. Lastly, we examined the results after excluding from our sample young women who were already pregnant at

the time of the 1980 interview and for whom educational and fertility expectations may have been directly influenced by the pregnancy.

Tables 8–10 show estimation results after modeling the propensity variable in the second-stage regression as a three-piece linear spline (bottom one-third, middle one-third, and top one-third of the propensity score distribution). As before, Table 8 shows estimates based on the propensity model without education and fertility expectations, and Tables 9 and 10 show results derived from the propensity model that controls for the expectations without nonlinear or interaction effects, and with nonlinear and interaction effects, respectively. The estimates confirm that the propensity variable is negatively related to educational and labor market achievement in all three ranges. However, the relationship between the propensity variable and subsequent educational and labor market outcomes appears to be nonlinear, with the negative gradient being the strongest in the bottom one-third of the propensity distribution, and the weakest in the top one-third of the distribution. While it appears that earnings and the likelihood of educational attainment decrease with higher propensity at a decreasing rate, further research is needed to examine possible explanations and draw conclusions from this observation. In the context of the present study, we focus on the robustness of the coefficients of the pregnancy variable. The estimates show that modeling the propensity as a spline rather than a single linear variable does not significantly change the estimates of the childbirth variable or any of the earlier findings.

TABLE 8 Non-Parametric Spline Propensity Score Estimates, Not Controlling for Expectations

	Outcomes					
	1982		**1992**			
	High School Dropout	**High School Dropout/ Delay**	**High School Diploma**	**Two-Year Degree or Higher**	**Four-Year Degree or Higher**	**Earnings**
Local estimates						
At low risk						
High school pregnancy	0.048 (2.80)***	0.078 (3.19)***	−0.049 (2.94)***	−0.150 (4.55)***	−0.193 (8.08)***	−2.452 (2.59)***
p-score spline						
Bottom third	0.412 (3.51)***	0.487 (3.90)***	−0.372 (3.14)***	−1.115 (7.28)***	−1.360 (10.99)***	−16.018 (3.73)***
Middle third	0.169 (4.52)***	0.0287 (5.37)***	−0.180 (4.96)***	−0.461 (5.56)***	−0.324 (4.23)***	−13.569 (5.70)***
Top third	0.018 (1.87)*	0.037 (2.31)**	−0.014 (1.57)	−0.020 (0.69)	−0.111 (3.36)***	−1.945 (2.29)**
At low-to-medium risk						
High school pregnancy	0.051 (2.93)***	0.085 (3.44)***	−0.052 (3.06)***	−0.149 (4.49)***	−0.188 (5.81)***	−2.440 (2.69)***
p-score spline						
Bottom third	0.444 (3.24)	0.506 (3.57)***	−0.402 (2.90)***	−1.103 (6.64)***	−1.316 (10.05)***	−15.984 (3.49)***
Middle third	0.182 (4.44)	0.304 (5.35)***	−0.196 (4.87)***	−0.471 (5.59)***	−0.324 (4.28)***	−13.639 (5.73)***
Top third	0.019 (2.10)**	0.036 (2.43)**	−0.015 (1.77)*	−0.017 (0.63)	−0.103 (3.61)***	−1.895 (2.55)**
At medium risk						
High school pregnancy	0.074 (2.18)**	0.133 (3.02)***	−0.075 (2.25)**	−0.140 (3.15)***	−0.138 (6.01)***	−2.537 (2.23)**

p-score spline						
Bottom third	0.740 (0.49)	0.593 (0.44)	−0.662 (0.43)	−0.784 (0.65)	−0.838 (1.12)	−15.056 (0.53)
Middle third	0.0286 (1.55)	0.452 (2.04)**	−0.318 (1.71)*	−0.544 (−2.30)**	−0.311 (1.91)*	−14.286 (2.50)**
Top third	0.031 (1.91)*	0.037 (1.63)	−0.026 (1.66)*	−0.004 (0.16)	−0.065 (2.62)***	−1.631 (2.29)**
At medium-to-high risk						
High school pregnancy	0.086 (1.98)**	0.157 (2.83)***	−0.088 (2.05)**	−0.152 (2.85)***	−0.136 (5.21)***	−3.231 (2.35)**
p-score spline						
Bottom third	0.801 (0.34)	0.596 (0.29)	−0.731 (0.30)	−0.796 (0.45)	−0.751 (0.71)	−13.041 (0.32)
Middle third	0.296 (1.06)	0.477 (1.47)	−0.322 (1.14)	−0.519 (1.56)	−0.348 (1.59)	−16.306 (2.04)**
Top third	0.034 (1.93)*	0.032 (1.33)	−0.031 (1.81)*	−0.010 (0.33)	−0.043 (1.76)*	−0.895 (1.22)
At high risk						
High school pregnancy	0.094 (1.98)**	0.169 (2.49)**	−0.096 (1.81)*	−0.160 (2.49)**	−0.136 (4.45)**	−3.541 (2.14)**
p-score spline						
Bottom third	0.832 (0.26)	0.596 (0.22)	−0.765 (0.23)	−0.797 (0.34)	−0.708 (0.52)	−11.988 (0.22)
Middle third	0.299 (0.80)	0.488 (1.14)	0.323 (1.39)	−0.512 (−1.19)	−0.371 (1.33)	−17.344 (1.68)*
Top third	0.036 (1.70)*	0.030 (1.07)	−0.033 (1.65)*	−0.010 (0.29)	−0.033 (1.22)	−0.587 (0.70)
Weighted average of local estimates						
High school pregnancy	0.053 (15.69)***	0.088 (20.22)***	−0.054 (17.64)***	−0.149 (22.94)***	−0.185 (25.48)***	−2.498 (15.82)***

Notes: Shown are marginal effects from a Probit regression. Bottom row is weighted average of local estimates. Absolute values of *t* -statistic for local estimates and of chi-square statistic for the weighted average are in parenthesis.

*<10%. **<5%. ***<1%.

TABLE 9 Non-Parametric Spline Propensity Score Estimates, with Controls for Expectations in the Propensity Model

	Outcomes					
	1982		**1992**			
	High School Dropout	**High School Dropout/ Delay**	**High School Diploma**	**Two-Year Degree or Higher**	**Four-Year Degree or Higher**	**Earnings**
Local estimates						
At low risk						
High school pregnancy	0.013 (1.07)	0.015 (0.83)	−0.016 (1.37)	−0.084 (2.57)**	−0.126 (4.26)***	−1.474 (1.57)
p-score spline						
Bottom third	0.419	0.763	−0.451	−1.764	−1.799	−29.309

	(2.72)***	(4.76)***	(2.68)***	(10.41)***	(14.49)***	(6.07)***
Middle third	0.195 (6.00)***	0.294 (6.88)***	−0.192 (6.02)***	−0.529 (8.20)***	−0.594 (10.06)***	−12.130 (6.35)***
Top third	0.032 (6.12)***	0.065 (7.12)***	−0.028 (5.61)***	−0.049 (2.58)**	−0.099 (3.87)***	−1.700 (3.18)**
At low-to-medium risk						
High school pregnancy	0.015 (1.14)	0.021 (1.05)	−0.019 (1.46)	−0.078 (2.36)**	−0.118 (4.05)***	−1.439 (1.55)
p-score spline						
Bottom third	0.454 (2.42)**	0.799 (4.26)***	−0.492 (2.28)**	−1.749 (9.33)***	−1.720 (12.99)***	−29.089 (5.51)***
Middle third	0.222 (5.84)***	0.320 (6.65)***	−0.219 (5.86)***	−0.542 (9.78)***	−0.579 (9.67)***	−12.420 (6.22)***
Top third	0.034 (6.34)***	0.066 (7.48)***	−0.029 (5.75)***	−0.045 (2.55)**	−0.092 (3.99)***	−1.573 (3.24)***
At medium risk						
High school pregnancy	0.037 (0.89)	0.07 (1.45)	−0.052 (1.26)	−0.043 (0.82)	−0.046 (1.83)*	−1.408 (1.01)
p-score spline						
Bottom third	0.482 (0.06)	0.50 (0.09)	−0.782 (0.08)	−1.002 (0.29)	−0.581 (0.46)	−17.197 (0.19)
Middle third	0.704 (1.44)	0.72 (1.59)	−0.718 (1.46)	−0.684 (−.98)**	−0.343 (2.27)**	−16.309 (1.73)*
Top third	0.057 (3.64)***	0.09 (4.59)***	−0.046 (3.01)***	−0.025 (1.12)	−0.036 (2.21)**	−0.898 (1.52)
At medium-to-high risk						
High school pregnancy	0.017 (0.33)	0.073 (1.39)	−0.042 (0.79)	−0.051 (0.76)	−0.045 (1.52)	−1.329 (0.75)
p-score spline						
Bottom third	0.468 (0.04)	0.502 (0.05)	−0.763 (0.05)	−0.978 (0.19)	−0.526 (0.30)	−16.110 (0.12)
Middle third	0.770 (1.01)	0.726 (1.13)	−0.792 (1.03)	−0.678 (1.36)	−0.316 (1.57)	−16.954 (1.23)
Top third	0.061 (3.24)***	0.091 (3.84)***	−0.047 (2.56)***	−0.023 (0.90)	−0.031 (1.71)*	−0.780 (1.15)
At high risk						
High school pregnancy	0.007 (0.11)	0.062 (0.92)	−0.036 (0.57)	−0.054 (0.67)	−0.044 (1.25)	−1.283 (0.58)
p-score spline						
Bottom third	0.483 (0.03)	0.444 (0.05)	−0.784 (0.04)	−0.970 (0.15)	−0.501 (0.23)	−15.992 (0.09)
Middle third	0.790 (0.76)	0.784 (1.13)	−0.813 (0.78)	0.672 (1.03)	−0.304 (1.20)	−17.008 (0.94)
Top third	0.064 (2.84)***	0.089 (3.84)***	−0.049 (2.23)**	0.023 (0.75)	−0.030 (1.38)	−0.778 (0.97)
Weighted average of local estimates						
High school pregnancy	0.015 (1.80)	0.023 (2.07)	−0.020 (3.36)*	−0.078 (6.26)**	−0.114 (9.96)***	−1.447 (2.35)

Notes: Shown are marginal effects from a Probit regression. Bottom row is weighted average of local estimates. Absolute values of t -statistic for local estimates and of chi-square statistic for the weighted average are in parenthesis.

∗<10%. ∗∗<5%. ∗∗∗<1%.

TABLE 10 Non-Parametric Spline Propensity Score Estimates, Controlling for Expectations in the Propensity Model with Interaction Effects

	Outcomes					
	1982		**1992**			
	High School Dropout	**High School Dropout/ Delay**	**High School Diploma**	**Two-Year Degree or Higher**	**Four-Year Degree or Higher**	**Earnings**
Local estimates						
At low risk						
High school pregnancy	0.016 (1.23)	0.027 (1.32)	0.019 (1.51)	−0.091 (2.67)∗∗∗	−0.120 (3.85)∗∗∗	−1.333 (1.38)
p-score spline						
Bottom third	0.704 (3.93)∗∗∗	0.941 (5.35)∗∗∗	−0.630(3.58)∗∗∗	−1.794 (8.78)∗∗∗	−1.997 (12.69)∗∗∗	−23.933 (4.10)∗∗∗
Middle third	0.116 (4.08)∗∗∗	0.198 (4.90)∗∗∗	−0.117 (4.21)∗∗∗	−0.378 (5.98)∗∗∗	−0.422 (7.32)∗∗∗	−10.037 (5.46)∗∗∗
Top third	0.028 (5.55)∗∗∗	0.051 (5.82)∗∗∗	−0.025 (5.11)∗∗∗	−0.047 (2.61)∗∗∗	−0.100 (4.47)∗∗∗	−1.884 (3.77)∗∗
At low-to-medium risk						
High school pregnancy	0.018 (1.29)	0.032 (1.50)	−0.021 (1.57)	−0.084 (2.48)**	−0.114 (3.75)∗∗∗	−1.190 (1.26)
p-score spline						
Bottom third	0.761 (3.49)∗∗∗	0.985 (4.78)∗∗∗	−0.683 (3.19)∗∗∗	−1.782 (7.87)∗∗∗	−1.913 (11.34)∗∗∗	−23.652 (3.71)∗∗∗
Middle third	0.136 (4.08)∗∗∗	0.218 (4.81)∗∗∗	−0.136 (4.16)∗∗∗	−0.387 (5.79)∗∗∗	−0.418 (7.11)∗∗∗	−10.339 (5.39)∗∗∗
Top third	0.029 (5.65)∗∗∗	0.051 (5.97)∗∗∗	−0.025 (5.23)∗∗∗	−0.044 (2.68)∗∗∗	−0.090 (4.54)∗∗∗	−1.777 (3.94)∗∗∗
At medium risk						
High school pregnancy	0.041 (1.03)	0.08 (1.59)	−0.048 (1.22)	−0.041 (0.80)	−0.057 (2.28)∗∗	−0.454 (0.37)
p-score spline						
Bottom third	0.311 (0.03)	0.826 (0.08)	−0.363 (0.03)	−1.053 (0.16)	−0.640 (0.22)	−5.39 (0.03)
Middle third	0.525 (1.12)	0.533 (1.15)	−0.473 (1.04)	−0.466 (1.19)	−0.313 (1.67)∗	−14.69 (1.54)
Top third	0.040 (2.79)∗∗∗	0.057 (3.13)∗∗∗	−0.037 (2.69)∗∗∗	−0.032 (1.50)	−0.029 (2.03)**	−1.239 (2.51)∗∗
At medium-to-high risk						
High school pregnancy	0.022 (0.46)	0.056 (0.90)	−0.035 (0.72)	−0.044 (0.71)	−0.059 (2.08)∗∗	−0.348 (0.24)
p-score spline						
Bottom third	0.256 (0.01)	0.823 (0.05)	−0.293 (0.02)	−1.007 (0.10)	−0.554 (0.14)	−1.145 (0.00)

Middle third	0.584 (0.80)	0.564 (0.80)	−0.537 (0.75)	−0.468 (0.83)	−0.300 (1.20)	−16.325 (1.213)
Top third	0.042 (2.75)***	0.061 (3.14)***	−0.038 (2.55)**	−0.029 (1.32)	−0.023 (1.53)	−0.966 (1.92)
At high risk						
High school pregnancy	0.010 (0.18)	0.041 (0.55)	−0.025 (0.43)	−0.045 (0.60)	−0.059 (1.79)*	−0.264 (0.15)
p-score spline						
Bottom third	0.251 (0.01)	0.839 (0.04)	−0.287 (0.01)	−0.983 (0.07)	−0.520 (0.10)	−0.224 (0.00)
Middle third	0.608 (0.61)	0.572 (0.61)	−0.561 (0.58)	−0.469 (0.64)	−0.291 (0.92)	−16.677 (0.95)
Top third	0.045 (2.50)**	0.064 (2.86)***	−0.040 (2.30)**	0.028 (1.11)	−0.021 (1.24)	−0.929 (1.63)
Weighted average of local estimates						
High school pregnancy	0.018 (2.10)	0.033 (1.62)	−0.004 (2.76)	−0.083 (6.71)***	−0.111 (11.01)***	−1.174 (2.70)

Notes: Shown are marginal effects from a Probit regression. Bottom row is weighted average of local estimates. Absolute values of *t* -statistic for local estimates and of chi-square statistic for the weighted average are in parenthesis.

*<10%. **<5%. ***<1%.

Secondly, we examine robustness of the estimates to the inclusion in the second-stage model of the variables used in the first-stage propensity estimation. As the propensity is the predicted likelihood of childbirth conditional on the variables used in its estimation, adding these variables along with the propensity score as controls in the second-stage estimation should not change the estimates of the effect of childbirth on subsequent outcomes. However, if the first-stage model is misspecified and poorly incorporates information contained in the variables into the propensity scores, adding these variables directly to the second-stage model could have significant implications for the robustness of the achievement and earnings gap estimates and thus the validity of conclusions than can be drawn from them. The results are displayed in Tables 11–13.[9] Adding the variables used in forming the propensity scores directly to the second-stage model makes the coefficients of the propensity statistically insignificant and their signs inconsistent both across models and in some cases between the estimation points within one set of estimates. However, the coefficients of the childbirth variable, both locally and the weighted averages, remain very similar in magnitude and significance to the set of estimates obtained based on the original specification of the models (Tables 5–7). Thus, the estimates are robust to alternative model specifications.

TABLE 11 Non-Parametric Propensity Score Estimates, Not Controlling for Expectations, Adjusting for Covariates Used in Predicting Propensity

	Outcomes					
	1982		**1992**			
	High School Dropout	**High School Dropout/ Delay**	**High School Diploma**	**Two-Year Degree or Higher**	**Four-Year Degree or Higher**	**Earnings**
Local estimates						
At low risk						

High school pregnancy	0.050 (2.96)***	0.081 (3.36)***	−0.052 (3.12)***	−0.151 (4.66)***	−0.187 (7.89)***	−2.477 (2.64)***
Propensity	−0.14 (0.60)	−0.014 (0.41)	0.019 (0.82)	0.098 (1.82)*	0.034 (0.67)	−1.516 (0.99)
At low-to-medium risk						
High school pregnancy	0.053 (3.08)***	0.087 (3.61)***	−0.054 (3.22)***	−0.150 (4.78)***	−0.183 (8.15)***	−2.471 (2.75)***
Propensity	−0.013 (0.55)	0.018 (0.54)	0.018 (0.77)	0.097 (1.91)*	0.028 (0.29)	−1.092 (0.77)
At medium risk						
High school pregnancy	0.075 (2.22)** (	0.134 3.10)**	−0.076 (2.29)**	−0.140 (3.20)***	−0.134 (5.85)***	−2.603 (2.30)**
Propensity	0.014 (0.22)	−0.036 (0.43)	−0.011 (0.18)	0.096 (0.97)	0.015 (0.21)	1.498 (0.62)
At medium-to-high risk						
High school pregnancy	0.085 (1.98)**	0.155 (2.86)***	−0.085 (2.03)**	−0.151 (2.85)***	−0.133 (5.06)***	−3.252 (2.36)**
Propensity	0.018 (0.25)	−0.059 (0.62)	−0.022 (0.32)	0.073 (0.67)	0.072 (0.93)	3.652 (1.38)
At high risk						
High school pregnancy	0.092 (1.73)*	0.166 (2.50)***	−0.091 (1.76)*	−0.157 (2.48)**	−0.133 (4.32)***	−3.535 (2.10)**
Propensity	0.023 (0.26)	−0.069 (0.60)	−0.028 (0.33)	0.074 (0.57)	0.095 (1.06)	4.509 (1.42)
Weight average of local estimates						
High school pregnancy	0.055 (16.91)***	0.090 (21.09)***	−0.056 (19.18)***	−0.150 (24.95)***	−0.180 (27.90)***	−2.527 (16.29)***

Notes: Shown are marginal effects from a Probit regression. Bottom row is weighted average of local estimates. Absolute values of t -statistic for local estimates and of chi-square statistic for the weighted average are in parenthesis.

*<10%. **<5%. ***<1%.

TABLE 12 Non-Parametric Propensity Score Estimates, Controlling for Expectations in the Propensity Score Model and Adjusting for Covariates Used in Predicting Propensity

	Outcomes					
	1982		**1992**			
	High School Dropout	**High School Dropout/ Delay**	**High School Diploma**	**Two-Year Degree or Higher**	**Four-Year Degree or Higher**	**Earnings**
Local estimates						
At low risk						
High school pregnancy	0.015 (1.30)	0.015 (0.85)	−0.018 (1.60)	−0.078 (2.50)**	−0.103 (3.65)***	−1.078 (1.16)
Propensity	0.010 (0.91)	0.028 (1.73)*	0.000 (0.08)	0.093 (3.45)***	0.054 (1.94)*	−1.759 (1.18)
At low-to-medium risk						
High school pregnancy	0.016 (1.35)	0.020 (1.09)	−0.020 (1.67)*	−0.073 (2.35)**	−0.098 (3.51)***	−1.074 (1.20)
Propensity	0.009	0.030	0.000	0.090	0.052	−1.407

	(0.76)	(1.80)∗	(0.07)	(3.44)∗∗∗	(1.97)∗∗	(1.01)
At medium risk						
High school pregnancy	0.028 (0.74)	0.066 (1.33)	0.042 (1.10)	−0.041 (0.81)	−0.041 (1.62)	−1.310 (1.14)
Propensity	−0.003 (0.05)	0.053 (0.70)	0.027 (0.42)	0.063 (0.86)	0.009 (0.21)	0.423 (0.18)
At medium-to-high risk						
High school pregnancy	0.008 (0.16)	0.053 (0.84)	−0.030 (0.60)	−0.045 (0.70)	0.041 (1.36)	−1.942 (1.38)
Propensity	0.011 (0.14)	0.045 (0.47)	0.013 (0.17)	0.052 (0.60)	−0.000 (0.01)	2.217 (0.85)
At high risk						
High school pregnancy	−0.001 (0.03)	0.047 (0.59)	−0.024 (0.38)	−0.047 (0.58)	0.040 (1.12)	−2.192 (1.26)
Propensity	0.022 (0.22)	0.045 (0.39)	0.002 (0.02)	0.048 (0.46)	−0.003 (0.06)	2.935 (0.93)
Weight average of local estimates						
High school pregnancy	0.016 (0.52)	0.022 (0.35)	−0.016 (2.73)	−0.072 (5.82)∗∗	−0.090 (8.41)∗∗∗	−1.139 (1.25)

Notes: Shown are marginal effects from a Probit regression. Bottom row is weighted average of local estimates. Absolute values of t -statistic for local estimates and of chi-square statistic for the weighted average are in parenthesis.

∗<10%. ∗∗<5%. ∗∗∗<1%.

TABLE 13 Non-Parametric Propensity Score Estimates, Controlling for Expectations in the Propensity Score Model with Nonlinear and Interaction Effects, and Adjusting for Covariates Used in Predicting Propensity

	Outcomes					
	1982		**1992**			
	High School Dropout	**High School Dropout/ Delay**	**High School Diploma**	**Two-Year Degree or Higher**	**Four-Year Degree or Higher**	**Earnings**
Local estimates						
At low risk						
High school pregnancy	0.019 (1.56)	0.028 (1.49)	−0.023 (1.87)∗	−0.088 (2.79)∗∗	−0.106 (3.69)∗∗∗	−1.355 (1.44)
Propensity	−0.011 (1.12)	−0.025 (1.75)∗	0.015 (1.46)	0.075 (3.11)∗∗∗	0.054 (2.26)∗∗	−0.025 (0.04)
At low-to-medium risk						
High school pregnancy	0.022 (1.66)∗	0.033 (1.69)∗	−0.026 (1.95)∗	−0.082 (2.61)∗∗∗	−0.100 (3.60)∗∗∗	−1.255 (1.36)
Propensity	−0.014 (1.26)	−0.025 (1.74)∗	0.016 (1.51)	0.071 (3.04)∗∗∗	0.054 (2.41)∗∗	0.030 (0.05)
At medium risk						
High school pregnancy	0.053 (1.36)	0.075 (1.58)	−0.056 (1.46)	−0.030 (0.61)	−0.044 (1.74)∗	−0.670 (0.52)
Propensity	−0.047 (0.80)	−0.043 (0.65)	0.039 (0.68)	0.043 (0.60)	0.039 (0.99)	1.103 (0.60)
At medium-to-high risk						
High school pregnancy	0.023 (0.49)	0.044 (0.76)	−0.032 (0.68)	−0.029 (0.46)	−0.047 (1.64)	−0.558 (0.33)

Propensity	−0.024 (0.36)	−0.017 (0.22)	0.027 (0.42)	0.038 (0.48)	0.017 (0.39)	0.697 (0.32)
At high risk						
High school pregnancy	0.009 (0.15)	0.027 (0.39)	−0.020 (0.34)	−0.026 (0.35)	−0.47 (1.39)	−0.468 (0.20)
Propensity	−0.015 (0.18)	−0.004 (0.05)	0.022 (0.27)	0.040 (0.42)	0.009 (0.19)	0.367 (0.13)
Weight average of local estimates						
High school pregnancy	0.022 (2.52)	0.033 (2.35)	−0.026 (2.73)	−0.079 (6.90)∗∗	−0.106 (10.12)∗∗∗	−1.234 (1.25)

Notes: Shown are marginal effects from a Probit regression. Bottom row is weighted average of local estimates. Absolute values of t -statistic for local estimates and of chi-square statistic for the weighted average are in parenthesis.

∗<10%. ∗∗<5%. ∗∗∗<1%.

TABLE 14 Non-Parametric Propensity Score Estimates, Not Controlling for Expectations, Excluding Pregnancies and Births Prior to 1980

	Outcomes					
	1982		**1992**			
	High School Dropout	**High School Dropout/ Delay**	**High School Diploma**	**Two-Year Degree or Higher**	**Four-Year Degree or Higher**	**Earnings**
Local estimates						
At low risk						
High school pregnancy	0.024 (1.89)∗	0.051 (2.32)∗∗	−0.024 (1.96)∗∗	−0.179 (4.24)∗∗∗	−0.249 (5.72)∗∗∗	−3.058 (2.77)∗∗∗
Propensity	0.034 (14.46)∗∗∗	0.078 (16.66)∗∗∗	−0.031 (14.05)∗∗∗	−0.223 (12.33)∗∗∗	−0.392 (14.53)∗∗∗	−6.020 (11.28)∗∗∗
At low-to-medium risk						
High school pregnancy	0.062 (2.29)∗∗	0.097 (2.78)∗∗∗	−0.063 (2.38)∗∗	−0.159 (4.74)∗∗∗	−0.114 (7.38)∗∗∗	−3.037 (2.87)∗∗∗
Propensity	0.084 (6.89)∗∗∗	0.140 (8.56)∗∗∗	−0.080 (6.77)∗∗∗	−0.194 (11.41)∗∗∗	−0.209 (25.01)∗∗∗	−5.47 (11.22)∗∗∗
At medium risk						
High school pregnancy	0.097 (1.98)∗∗	0.142 (2.59)∗∗	−0.096 (2.00)**	−0.136 (2.85)∗∗∗	−0.075 (4.01)∗∗∗	−3.153 (2.42)**
Propensity	0.067 (2.65)∗∗∗	0.077 (2.79)∗∗∗	−0.062 (2.50)∗∗	−0.046 (1.89)∗	−0.063 (8.49)∗∗∗	−2.575 (4.17)∗∗∗
At medium-to-high risk						
High school pregnancy	0.138 (2.03)∗∗	0.179 (2.61)∗∗∗	−0.140 (2.07)∗∗	−0.142 (2.49)∗∗	−0.058 (2.58)∗∗∗	−3.978 (2.55)∗∗
Propensity	0.072 (2.00)∗∗	0.062 (1.89)∗	−0.068 (1.92)∗	−0.033 (1.34)	−0.036 (9.09)∗∗∗	−1.667 (2.50)∗∗
At high risk						
High school pregnancy	0.167 (1.88)∗	0.198 (2.37)∗∗	−0.171 (1.92)∗	−0.143 (2.07)∗∗	−0.046 (1.57)∗	−4.358 (2.32)∗∗
Propensity	0.078 (1.56)	0.056 (1.38)	−0.075 (1.50)	−0.028 (1.01)	−0.024 (7.86)∗∗	−1.314 (1.69)∗

Weight average of local estimates						
High school pregnancy	0.050 (10.86)***	0.082 (12.33)***	−0.051 (10.33)***	−0.166 (20.29)***	−0.173 (18.88)***	−3.110 (20.96)***

Notes: Shown are marginal effects from a Probit regression. Bottom row is weighted average of local estimates. Absolute values of t -statistic for local estimates and of chi-square statistic for the weighted average are in parenthesis.

*<10%. **<5%. ***<1%.

Our last robustness check is motivated by the fact that almost 80 young women in our sample had already had a child or were pregnant at the time of the 1980 interview. As these women may have adjusted their educational and fertility expectations as a result of childbirth, including them in the earlier analysis may have overstated the role of expectations in explaining the achievement and earnings gaps associated with adolescent motherhood. The analysis of the women's educational and fertility expectations in 8th grade (recorded retrospectively in the 1980 sophomore interview) showed high correlations between the pre- and post-childbirth sets of expectations (correlation coefficients between expectations in 8th and 10th grade were .82–.94). Nevertheless, in order to have confidence in our conclusions about the role of expectations, it is necessary to examine whether the results are robust to excluding women with potentially contaminated expectations.

The results of the analysis after excluding women who were pregnant or had a child prior to the 1980 interview are shown in Tables 14–16. In all tables, the propensity variable is modeled as a single linear variable as it was in the original specification of the model. A comparison of estimates between Tables 5 and 14 shows that, overall, the impact of childbirth on subsequent educational outcomes and earnings is slightly smaller when the women who were pregnant or had a child prior to 1980 are excluded. This is because of the fact that the women who became pregnant in 10th grade had systematically lower educational achievement and earnings compared to young mothers who became pregnant in grades 11 or 12; therefore excluding them from the sample makes the statistical effect of childbirth on subsequent outcomes appear slightly smaller. Having fewer young mothers in the sample also increases the standard errors of the coefficients, further reducing their statistical significance. Nevertheless, controlling for educational and fertility expectations produces a virtually identical magnitude reduction in the estimates of the effect of childbirth on subsequent educational outcomes and earnings. For example, the average gap in achieving a 4-year degree or higher decreases by 8.7% points (from 17.3% in Table 14), compared to a 9% point reduction between Tables 5 and 7. As a result, all of the educational achievement gaps in Table 16 are smaller in magnitude than the corresponding estimates in Table 7. While the earnings gap appears higher (−$1,726 vs. −$988 in Tables 14 and 5, respectively), it is statistically insignificant overall as well as all of its local estimates. Thus, we conclude that our finding of the role of expectations is robust to excluding women for whom expectations may have been contaminated by prior pregnancy.

TABLE 15 Non-Parametric Propensity Score Estimates, Controlling for Expectations in the Propensity Score Model, Excluding Pregnancies and Births Prior to 1980

	Outcomes					
	1982		**1992**			
	High School Dropout	**High School Dropout/ Delay**	**High School Diploma**	**Two-Year Degree or Higher**	**Four-Year Degree or Higher**	**Earnings**
Local estimates						
At low risk						

High school pregnancy	0.001 (0.20)	0.002 (0.18)	−0.002 (0.43)	−0.090 (2.12)**	−0.166 (2.83)***	−1.678 (1.08)
Propensity	0.029 (14.62)***	0.075 (22.97)***	−0.026 (13.93)***	−0.256 (18.76)***	−0.621 (22.77)***	−5.247 (13.80)***
At low-to-medium risk						
High school pregnancy	0.010 (0.32)	0.015 (0.37)	−0.019 (0.59)	−0.058 (2.08)**	−0.005 (2.87)***	−1.645 (1.54)
Propensity	0.151 (8.65)***	0.225 (12.39)***	−0.137 (8.19)***	−0.182 (30.32)***	−0.029 (5.50)***	−4.731 (13.16)***
At medium risk						
High school pregnancy	0.055 (0.78)	0.075 (1.12)	−0.074 (1.08)	−0.043 (0.87)	−0.007 (0.96)	−2.202 (1.39)
Propensity	0.124 (3.68)***	0.140 (5.47)***	−0.102 (3.11)***	−0.041 (2.63)***	−0.12 (2.84)***	−1.513 (2.59)***
At medium-to-high risk						
High school pregnancy	0.042 (0.46)	0.057 (0.83)	−0.081 (0.85)	−0.047 (0.86)	−0.001 (0.52)	−2.317 (1.14)
Propensity	0.139 (5.25)***	0.114 (12.74)***	−0.120 (3.33)***	−0.032 (2.96)***	−0.002 (0.83)	−1.316 (1.90)*
At high risk						
High school pregnancy	0.023 (0.29)	0.030 (0.59)	−0.067 (0.65)	−0.040 (0.70)	−0.000 (0.27)	−2.362 (0.94)
Propensity	0.102 (6.06)***	0.065 (3.06)***	−0.109 (9.71)***	−0.025 (5.71)***	−0.000 (0.35)	−1.266 (1.53)
Weight average of local estimates						
High school pregnancy	0.009 (0.16)	0.014 (0.27)	−0.017 (0.58)	−0.072 (3.74)**	−0.081 (3.85)**	−1.726 (2.78)

Notes: Shown are marginal effects from a Probit regression. Bottom row is weighted average of local estimates. Absolute values of *t* -statistic for local estimates and of chi-square statistic for the weighted average are in parenthesis.

*<10%. **<5%. ***<1%.

TABLE 16 Non-Parametric Propensity Score Estimates, Controlling for Expectations in the Propensity Score Model with Interaction Effects, Excluding Pregnancies and Births Prior to 1980

	Outcomes					
	1982		**1992**			
	High School Dropout	**High School Dropout/ Delay**	**High School Diploma**	**Two-Year Degree or Higher**	**Four-Year Degree or Higher**	**Earnings**
Local estimates						
At low risk						
High school pregnancy	0.004 (0.47)	0.014 (0.85)	−0.005 (0.69)	−0.108 (2.52)**	−0.159 (2.99)***	−1.845 (1.67)*
Propensity	0.029 (14.97)***	0.068 (19.06)***	−0.026 (14.26)***	−0.195 (14.83)***	−0.429 (18.43)***	−4.417 (12.41)***
At low-to-medium risk						

High school pregnancy	0.017 (0.56)	0.039 (1.00)	−0.024 (0.80)	−0.079 (2.46)**	−0.021 (3.27)***	−1.669 (1.54)
Propensity	0.107 (7.84)***	0.155 (10.20)***	−0.097 (7.43)***	−0.153 (18.36)***	−0.071 (11.56)***	−4.04 (12.15)***
At medium risk						
High school pregnancy	0.052 (0.84)	0.078 (1.20)	−0.059 (0.98)	−0.041 (0.81)	−0.021 (1.36)	−1.119 (0.81)
Propensity	0.079 (2.85)***	0.087 (3.36)***	−0.074 (2.71)***	−0.041 (2.54)**	−0.019 (7.34)***	−1.658 (3.37)***
At medium-to-high risk						
High school pregnancy	0.055 (0.62)	−0.060 (0.75)	−0.073 (0.84)	−0.046 (0.83)	−0.014 (0.91)	−1.223 (0.73)
Propensity	0.104 (3.16)***	0.095 (4.07)***	−0.098 (2.82)***	−0.033 (2.75)***	−0.009 (2.99)***	−1.370 (2.59)***
At high risk						
High school pregnancy	0.045 (0.45)	0.039 (0.49)	−0.069 (0.65)	−0.041 (0.71)	−0.007 (0.54)	−1.239 (0.61)
Propensity	0.106 (5.47)***	0.083 (9.47)***	−0.104 (4.06)***	−0.027 (4.05)***	−0.004 (1.09)	−1.308 (2.15)**
Weight average of local estimates						
High school pregnancy	0.014 (0.54)	0.026 (1.31)	−0.019 (1.06)	−0.089 (5.37)**	−0.086 (5.51)**	−1.701 (2.92)

VI. CONCLUSION

This study uses a locally weighted propensity score regression to examine the importance of educational and fertility expectations in explaining the achievement gap of adolescent mothers along a continuum of socioeconomic environments. We find that while the economic disadvantage associated with having a child or becoming pregnant in high school appears to be particularly large among young women of poor socioeconomic status, this disadvantage is as a result of preexisting differences in the educational and fertility expectations and is not as a result of a diminished capacity of young mothers of low socioeconomic status to cope with an unplanned early pregnancy.

These findings suggest that governmental programs that provide childcare and other financial assistance to young mothers with low means may not be enough to improve their long-term well-being and financial security. Nor are these programs likely to be very successful in enhancing social mobility and reducing income inequality. In fact, as most of the negative association between early pregnancy and low educational achievement and earnings appears to be a realization of women's a priori educational and fertility plans, programs that reduce the cost of an early childbirth can increase the incidence of early pregnancies and further worsen educational attainment and earnings of young mothers.

This study underscores the role of educational and fertility expectations as facilitators of a young woman's future well-being, and as drivers of the poverty cycle and income inequality. Thus, we advocate public policy efforts focused on encouraging young women to approach parenthood responsibly and seek schooling as a means of becoming a successful adult. An example of one such policy can be found in [18] who finds that students from low-income families tend to underestimate the return to schooling and that informing them about the true return to schooling can make the students significantly less likely to drop out of high school. Our study further underscores that such efforts ought to be undertaken early on, as by the sophomore year of high

school many young women already have well-formed expectations of their future that strongly correlate with their subsequent outcomes.

Future research identifying factors that have the strongest influence on young women's educational and fertility expectations may reveal potential policy levers that can be used to lower the incidence of adolescent childbirth, improve the educational achievement and labor market outcomes of adolescent mothers, and help break the poverty cycle.

Footnotes

1 In compliance with National Center for Educational Statistics restricted data rounding rules, all sample size numbers are rounded to the nearest 10.

2 1991 is the last full year of earnings data in the HS&B survey.

3 The non-parametric regression technique used in this study (Fan 1992) allows flexibility in the choice of the estimation grid. While an equally spaced grid, such as the one used in this study, is offered, other types of estimation grids (e.g., distribution-based grids) are allowed. While both equally spaced and distribution-based grids were examined in this study, the results of the equally spaced grid proved to be more informative because of the distributional properties of the propensity score variable (discussed in more detail later). The results obtained using a distribution-based grid (deciles and quintiles) are available from the authors upon request.

4 Note that although many of the variables used in estimating the propensity in the first step (e.g., socioeconomic and demographic characteristics, expectations) are likely to have a direct impact on subsequent educational and labor market outcomes, they are not included as controls in estimating the impact of childbirth on subsequent outcomes. The reason is that conditional on the predicted value of the propensity, the actual childbirth variable is orthogonal to the set of covariates that were used to predict the propensity; therefore, not including them in the second-step model will not bias the coefficient of the childbirth variable.

5 The estimation was performed using the seemingly unrelated estimation technique [suest] in Stata 11.0.

6 Propensity-based methods have been found to be less biased than both linear regression and fixed-effect estimation in some nonexperimental settings (Dehejia and Wahba 1999).

7 In Table 1, 15.38% of young mothers dropped out, compared to only 4.8% of women who delayed childbirth; the difference between the two groups is 10.58%.

8 Note that since Stata 11.0 does not allow bootstrapping standard errors in weighted models, no correction is made for the first-stage estimation error. Therefore the standard errors of the coefficients of the propensity are likely to be understated.

9 Owing to space considerations, the coefficients of the first-stage covariates in the second-stage model are not shown in the tables. They are available upon request.

[*]*This study has greatly benefited from insightful comments of many people, including Brian Brush, David Clark, Todd Elder, and Roger Koenker. I am especially grateful to Kevin Hallock and to Darren Lubotsky for extraordinary feedback on an earlier draft. I also wish to thank two anonymous referees for this journal, as well as seminar participants who were in attendance when this article was presented at the University of Chicago, University of Illinois, and Marquette University.*

REFERENCES

Anderseck, K., and H. Günter. "The Validity of Subjective Information as a Formal Concept Applied to Empirical Analysis." *Theory and Decision*, 5, 1974, 23 – 43.

Attanasio, O., and K. Kaufmann. "Educational Choices, Subjective Expectations, and Credit Constraints." *NBER Working Paper 15087*, Orazio Attanasio and Katja Kaufmann, 2009.

Bloom, D. E., and J. Trussel. "What Are the Determinants of Delayed Childbearing and Permanent Childlessness in the United States?" *Demography*, 21, 1984, 591 – 611.

Bornstein, M. H. *Handbook of Parenting: Status and Social Conditions of Parenting*. Mahwah, NJ: Lawrence Erlbaum Associates, 1995.

Bronars, S. G., and J. Grogger. "The Economic Consequences of Unwed Motherhood: Using Twin Births as a Natural Experiment." *The American Economic Review*, 1994, 1141 – 56.

Corijn, M., A. C. Liefbroer, and J. D. Gierveld. "It Takes Two to Tango, Doesn't It? The Influence of Couple Characteristics on the Timing of the Birth of the First Child." *Journal of Marriage and Family*, 58, 1996, 117 – 26.

Dehejia, R. H., and S. Wahba. "Causal Effects in Non-experimental Studies: Re-evaluating the Evaluation of Training Programs." *Journal of the American Statistical Association*, 94, 1999, 1053 – 62.

Delavande A., and H. Kohler. "Subjective Expectations in the Context of HIV/AIDS in Malawi." *Demographic Research*, 20, 2009, 817 – 74.

Dominitz, J., and C. F. Manski. "Eliciting Student Expectations of the Returns to Schooling." *The Journal of Human Resources*, 31, 1996, 1 – 26.

Fan, J. "Design-Adaptive Nonparametric Regression." *Journal of American Statistical Association*, 87, 1992, 998 – 1004.

Finer, L., and S. Henshaw. *Estimates of U.S. Abortion Incidence 2001–2003*. New York: Guttmacher Institute, 2006.

Geronimous, A. T., and S. Korenman. "The Socioeconomic Consequences of Teen Childbearing Reconsidered." *The Quarterly Journal of Economics*, 107, 1992, 1187 – 214.

Gully, H. "Suddenly, Teen Pregnancy is Cool?" Accessed January 17, 2008. http://www.macleans.ca/culture/lifestyle/article.jsp?content=20080117%5f99497%5f99497.

Ham, J. C., X. Li, and P. B. Reagan. "Propensity Score Matching, a Distance-Based Measure of Migration, and the Wage Growth of Young Men." *Journal of Economic Literature*, 2003, 315 – 31.

Hamilton, B. E., J. A. Martin, and S. J. Ventura. "Births: Preliminary Data for 2006," *National Vital Statistics Reports*, vol. 56, no. 7. Hyattsville, MD: National Center for Health Statistics, 2007.

Hotz, V. J., C. H. Mullin, and S. Sanders. "Bounding Causal Effects Using Data from a Contaminated Natural Experiment: Analyzing the Effects of Teenage Childbearing." *Review of Economic Studies*, 64, 1997, 575 – 603.

Hotz, V. J., S. McElroy, and S. Sanders. "Teenage Childbearing and Its Life Cycle Consequences: Exploiting a Natural Experiment." *Journal of Human Resources*, 40, 2005, 683 – 715.

Jensen, R. "The (Perceived) Returns to Education and the Demand for Schooling." *Quarterly Journal of Economics*, 125 (2), 2010, 515 – 48.

Juster, T. "Consumer Buying Intentions and Purchase Probability: An Experiment in Survey Design." *Journal of the American Statistical Association*, 61, 1966, 658 – 96.

Kaufmann, K. "Understanding the Income Gradient in College Attendance in Mexico: The Role of Heterogeneity in Expected Returns to College." *Mimeo*, Stanford University, 2009.

Kingsbury, K. "Pregnancy Boom at Gloucester High." *Time Magazine*. June 18, 2008. http://www.time.com/time/world/article/0,8599,1815845,00.html.

Kravdal, O. "The Importance of Economic Activity, Economic Potential and Economic Resources for the Timing of First Births in Norway." *Population Studies*, 48, 1994, 249 – 67.

Manski, C. "Measuring Expectations." *Econometrica*, 72, 2004, 1329 – 76.

Marini, M. M. "Women's Educational Attainment and the Timing of Entry into Parenthood." *American Sociological Review*, 49, 1984, 491 – 511.

McCrate, E. "Expectations of Adult Wages and Teenage Childbearing." *International Review of Applied Economics*, 6, 1992, 309 – 28.

McKenzie, D., G. Gibson, and S. Stillman. "A Land of Milk and Honey with Streets Paved with Gold: Do Emigrants Have Over-Optimistic Expectations about Incomes Abroad?" *World Bank Policy Research Working Paper* No. 4141, 2007.

Nguyen, T. "Information, Role Models and Perceived Returns to Education: Experimental Evidence from Madagascar." *Mimeo*, MIT, 2008.

Rosenbaum, P. R., and D. B. Rubin "The central role of the Propensity Score in Observational Studies for Causal Effect." *Biometrika*, 70, 1983, 41 – 55.

Trent, K. "Family Context and Adolescents' Fertility Expectations." *Youth and Society*, 96, 1994, 118 – 37.

Trusty, J. "Family Influences on Educational Expectations of Late Adolescents." *The Journal of Educational Research*, 91, 1998, 260 – 70.

Waite, L. J., and K. A. Moore. "The Impact of an Early First Birth on Young Women's Educational Attainment." *Social Forces* 57, 1978, 845 – 65.

Zellner, A. "An Efficient Method of Estimating Seemingly Unrelated Regression Equations and Tests for Aggregation Bias." *Journal of the American Statistical Association*, 57, 1962, 348 – 68.